Automated Axon Segmentation from Highly Noisy Microscopic Videos

John Bowler, Rogerio Feris, Liangliang Cao, Jun Wang, Mo Zhou
Columbia University
116th St & Broadway, New York, NY 10027
jcb2238@columbia.edu, rsferis@us.ibm.com, liangliang.cao@us.ibm.com,
j.wang@alibaba-inc.com, mz2417@columbia.edu

Abstract

We present a novel method for automated segmentation of axons in extremely noisy videos obtained via two-photon microscopy in awake mice. We formulate segmentation as a pixel-wise classification problem in which a pixel is classified into "axon" or "non-axon" based on its feature vector. In order to deal with high levels of noise, the features of our classifier are derived from spatio-temporal Independent Component Analysis (stICA) which effectively isolates noise from signal components while leveraging temporal coherence from the video. We fit parametric models to represent the distribution of the extracted features and apply a probabilistic classifier over stICA components to determine the label of each pixel. Finally, we show compelling qualitative and quantitative results from very challenging two-photon microscopic, demonstrating the usefulness of our approach. An example time-series of two-photon images with our automated ROI extraction overlayed is available with the supplemental materials.

1. Introduction

Two-photon microscopy has emerged as an important tool for *in vivo* measuring the activity of large populations of neurons and sub-cellular neuronal compartments, such as axons and dendrites. By imaging genetically encoded calcium indicators [18] expressed in neurons of awake behaving animals, neuroscienctists are able to relate the activitiy of neuronal populations – reflected by changes in fluorescence intensity – to behavior and learning [7]. Detecting and segmenting these neuronal structures to identify regions-of-interest (ROI) for subsequent analysis is a fundamental task in this process, which is especially challenging in the case of small and irregularly shaped sub-cellular compartments such as axons and dendrites. Accurate automated solutions to the problem are necessary, as manual approaches are inefficient and inconsistent from one researcher to the next.

In this paper we address the problem of automating the segmentation of axons in two-photon calcium imaging data. A unique challenge to accomplish this goal dealing with extremely noisy data inherent in the two-photon imaging process, as shown in Figure 1. In fact, off-the-shelf segmentation algorithms proposed in the computer vision community [15, 12, 5] are clearly not suitable to process this kind of images. Also, standard image denoising techniques [13, 1] are generally designed to handle limited noise levels instead of extremely noisy frames as in our application.

As a novel contribution, we propose a supervised classification method for automated axon segmentation with well-crafted features suited to handle highly noisy two-photon microscopy. For features, we consider intensity and gradient measurements from component images obtained by applying dimensionality reduction followed by spatio-temporal independent component analysis (stICA) to the input video, which helps isolating noise from neuronal activity signals by separating areas of the images whose function is correlated in both space and time [17, 11]. Then, we model the distribution of features and design a probabilistic classifier to determine "axon" and "non-axon" pixels from stICA components. Finally, we apply connected component analysis and use a contour tracing algorithm to determine axon boundaries.

The rest of the paper is organized as follows. Section 2 provides an overview of related work. In Section 3 we describe the details of our proposed approach for segmentation of axons in two-photon microscopy videos. Section 4 covers our experimental analysis, including qualitative and quantitative results, as well as comparison with other methods. Finally, Section 5 concludes the paper.

2. Literature Review

The recent advances in high-resolution and rapid microscopy techniques provide scientists the cell-level visual perception to understand the underlying mechanism of complex biological systems. Due to the high-volume of the generated data, manual annotation and processing

978-1-4799-6683-7/15 $31.00 © 2015 IEEE
DOI 10.1109/WACV.2015.126

is infeasible for large scale morphological analysis of axons needed in real applications. Hence, interests in automatic segmentation techniques for microscopy images have been rapidly expanding since the segmentation results can be used further for various quantitative analysis including tracking cellular dynamics [16], identifying cellular types [20, 22, 19] and extracting axon morphology [8, 14]. As surveyed in [10], many techniques have been developed for segmenting various types of cell images in the past few decades. Most of the existing approaches rely heavily on several heuristic segmentation strategies, including intensity thresholding, morphological filtering and region accumulation. In addition, previous studies focus on segmenting cells from static microscopy images and tend to be highly sensitive to image noise. Therefore, these static approaches have limited capacity to explore the spatio-temporal information of dynamic signals from axons, neither can they handle high levels of imaging noise well.

Furthermore, some new approaches were specifically designed to extract the geometric structures of axons from a sequence of microscopy images [2, 14, 21]. For instance, the repulsive force based snake model is used to connect axonal profiles segmented from each image frame to reconstruct the 3D structure [2]. Additionally, a dynamic local tracing approach has been designed to retrieve the 3D curvilinear structure of axons [21]. However, in order to be successful for our application the segmentation must succeed while accounting for both extremely noisy data as well as the possibility of overlapping segments. Examples of two-photon microscopy imaging data are show in Figure 1 (a) and the supplemental materials. To address these challenges, we propose a novel segmentation procedure that explores stICA and pixel level classification as illustrated in Figure 1. The stICA technique has been used to process functional magnetic resonance imaging data [17] and also to extract cell locations and their dynamics with minimum supervision from calcium imaging data [11]. Pixel classification has been used for automated segmetation of ultrasound images in [23]. These methods are combined to develop a fully automated axon segmentation approach which perfroms well and is capable of extracting ROIs even when presented with highly noisy data.

3. Technical Approach

Our approach focuses on two steps, outlined in Figure 1. In the first step, stICA is used to identify the components of the data which statistically have maximal independance in both space and time. The second step is then to extract the ROIs from each of the stICA components through the application of a pixelwise classifier on each of the extracted components.

3.1. Preprocessing

Raw video files are acquired from awake head-fixed mice performing tasks on a treadmill while positioned under a two-photon microscope (see [9] for methods). An example of the input data is shown in Figure 1 (a). The resulting images require motion correction both to account for the motion of the animal on the treadmill as well as the scanning techniques implemented by the microscope. Therefore, prior to extracting components from the data, the time series is preprocessed using established techniques to correct for motion artifacts within and between frames using an established Hidden Markov Model (HMM) based technique [4, 9].

3.2. Spatio-Temporal Component Analysis

We are interested in spatio-temporal components which can be used to help identify axon segments and remove noise. One method which has been applied to find the components is PCA. Prior to conducting PCA, each video is represented as a matrix $\mathbf{M} \in \mathbb{R}^{N_x \times N_T}$, where N_x is the number of pixels in each frame and N_t is the number of frames in the time series. Applying PCA to a video of the data picks out the dimensions along which the intensity of the pixels have the maximum variance. By performing PCA on the time-series and maintaining only the first k principal components (PCs), dimensions containing low amounts of information are removed resulting in a significant decrease in noise. Qualitative results confirm that this method generates components with less noise in them (see supplemental Materials). Note that it has been reported that the first few spatial PCs tended to be dominated by high variance noise sources [11]. However, in our experiments, we maintained all of the first PCs because they often contained useful axon information (see supplemental materials).

Although PCA helps in identifying the image regions with active axons, single PCs often contain overlapping segments of distinct axons. To separate the extracted axons, we consider applying ICA that is developed for separating linearly mixed signals. ICA finds the components of the signal which are statistically independent, rather than simply uncorrelated. Applying ICA to the extracted PCs is a process that takes into consideration both the *spatial* and *temporal* coefficients and tends to generate sparser components than those found through PCA alone. In order to implement the spatial and temporal ICA, the temporal coefficients from the PCA are concatenated to the spatial ones as suggested in [11]. Mathematically, we define the following spatio-temporal variable y_i and the corresponding ICA components z_i:

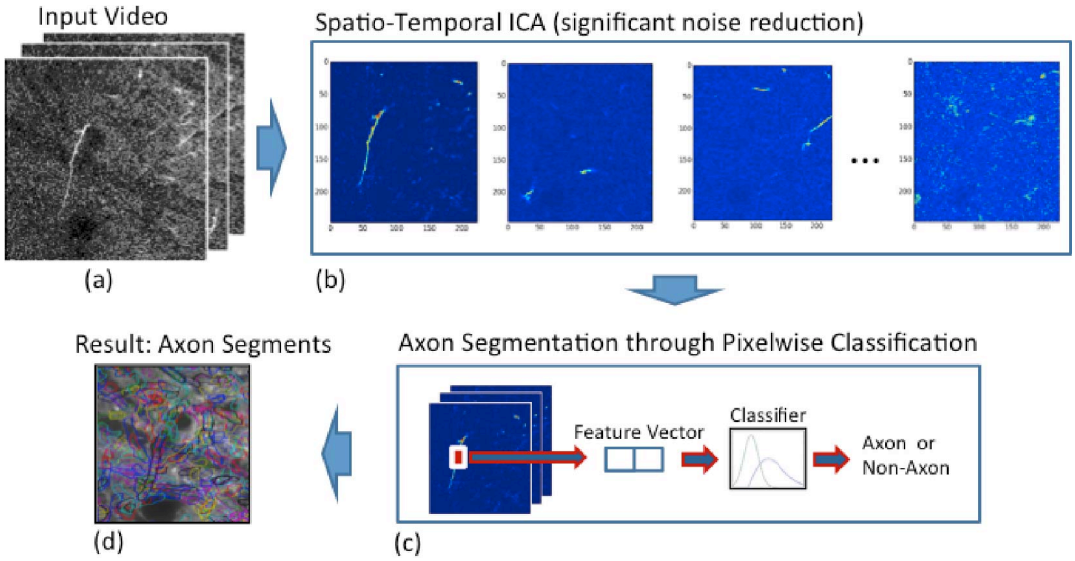

Figure 1. (a) Time sequences of two-photon images of calcium signals from axons of hippocampal neurons in awake mice performing behavior and learning tasks. (b) Extracted component images reflecting regions of activity which are maximally indepentdant in both space and time, through the application of the stICA methods. (c) Pixelwise classifiers are trained and distributions for two features are extracted, describing the stICA pixel intensity and local gradients. (d) Evaluating each pixel and extracting those most likely to be representing neuronal axons allows boundaries to be defined which separate the different functional units in the time sequence.

$$y_i = \begin{cases} \mu U_{ki} & i \leq N_x \\ (1-\mu)\, V_{ki} & N_x < i \leq (N_x + N_t) \end{cases} \qquad (1)$$
$$z_i^k = \sum_j \mathbf{W}_{i,j}^{(n)} y^{(j)},$$

where k is the number of components used for the analysis, U corresponds to the spatial PCA component matrix with dimensions N_x by k PCs and V corresponds to the N_t by k temporal PCA component matrix. Here μ is a weighting parameter to balance the tradeoff between the spatial and temporal information. ICA is performed on y_i to extract the independent components z_i. In our experiments, we use the FastICA algorithm [6] to compute the weighting matrix $\mathbf{W}$, as well as the independent components z_i. Figure 1 (b) shows examples of components extracted by stICA. Since stICA components are also temporally independent active axon segments with overlapping pixels in the average of the time series appear in separate stICA components if they are active at different times. This is advantageous since it allows the axons which overlap to be identified in separate segments. The stICA components result in a much sparser representation of the axons which simplifies the segmentation problem.

3.3. Axon Segmentation by Pixelwise Classification

In order to extract the ROIs (axon segments) from the stICA components, individual pixels from each component are classified as "axon" or "non-axon", as to whether they belong to regions of interest or not. Our classifier operates at the pixel-level and relies on two descriptive features derived from the stICA component images The first feature is the relative weighting of the individual pixels in each stICA components. The second is the sum of the magnitudes of the gradients in a small area surrounding each pixel.

Prior to extracting this feature vector, we subtract the mean of every component and normalize the absolute values to be in the range between 0 and 1. Subtraction and normalization ensure that pixel weights in axon regions are normalized across components and that the relative importance of each pixel is maintained. This normalization makes it possible to train likelihood distributions using training datasets and then run the classifier on any stICA components.

In the training phase, we manually label a set of positive samples (pixels belonging to ROI/axon regions) and negative samples (pixels outside ROI/axon regions) from component images obtained by stICA. We use these training samples to model the distribution of our features.

Figure 2 (a) and (b) illustrate the binning of our first descriptive feature (stICA pixel intensity) for ROI pixels, showing that they fit approximately to a Rayleigh distribution. The Rayleigh distribution is chosen, as in [23], for the ROI pixels due to it's shape and fit to the histograms of the data. We also show that the non-ROI pixels fit well to a normal distribution. Similar results are achieved by plotting these distributions for several of the components.

Our second descriptive feature used to classify pixels is

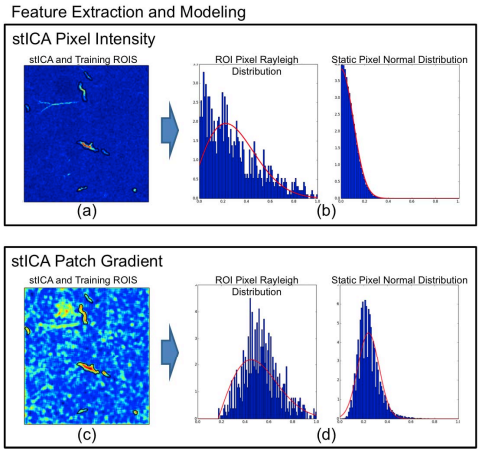

Figure 2. (a) and (c) Training axons are selected from stICA components, and used to generate distributions which describe the pixel intensity and the negative sum of local gradients. (b) and (d) Axon pixels are fit to a Rayleigh distributions while the noisy background pixels are fit to normal distributions

the evaluation of the patch-gradient for each of the stICA pixels. In order to get an accurate representation of whether a pixel is sitting on an area of random intensities, or near an area of similarly weighted ROI pixels, a sum of gradients is conducted in a small area around each pixel. For this to work on a small scale, the image is converted to a binary array first based on a threshold on the standard deviation of the pixels in each component. Then the gradients are summed up around the local neighborhood of each pixel.

Training on ROIs from several stICA components is taken to ensure the distributions generalize well when applied to components on which they have not been trained.

Figure 2 (c) and (d) illustrate the extraction of patch-gradient features. The ROI patch gradients are fit to another Rayleigh distribution while the non-Axon fits a normal distribution.

During the testing phase, the following steps are performed to identify axon/ROI pixels. First, the input video is pre-processed and stICA is conducted as described in the previous section. Then, for each pixel of each component, we extract our two-dimensional feature vector (stICA pixel intensity and patch-gradient) and use the parametric feature distributions computed at training time to measure the likelihood of each pixel to be part of an axon or part of the background. For every pixel both the "axon" likelihood and the "non-axon" likelihoods are evaluated. The pixel is assigned a category based on which likelihood is higher. An example of this training step for a single stICA component is shown in Figure 3. The previously trained distributions for both "axon" and "non-axon" classes are shown in (a) and (c) while the weights for each pixel and resulting class decision are indicated in Figure 3 (b).

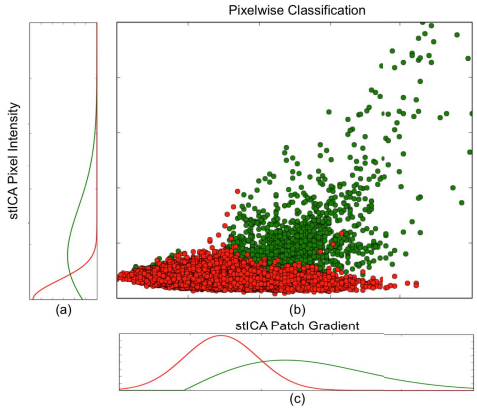

Figure 3. Distributions to model the likelihood that an individual pixel is either an ROI or background noise. (a) For the stICA intensity feature, a Rayleigh distribution models the ROI pixel intensity and a normal distribution centered near zero models the background. (c) Similar distributions were found which model the stICA patch gradient feature. (b) Both features are extracted for every pixel in a single stICA component and pixels are classified as to whether they are more likely to reside within an ROI, shown in green, or more likely to reside in background, shown in red.

3.4. Spatial Segmentation

After the classification step each of the stICA components has been translated into a binary mask of pixels labeled as belonging to the "axon" class. These masks are then separated completely along any discontiguous units into individual ROIs. This can be easily implemented via connected component analysis, i.e. a recursive search for adjacent non-zero pixel values. Note that this process is applied to each stICA component separately, where the axons are not overlapped. No additional pre-processing based on morphological operations was required.

ROI masks which contain very few pixels are discarded at this point as they are not large enough to perform useful analysis and most likely arise from background noise. We then apply a clockwise traversal of the ROI boundary to connect edge points with their neighbors in an ROI region. This optional step smooths the ROI boundary and facilitates future analysis.

4. Experimental Results

Our classification system was run on datasets of 3500 frames of two-photon data taken from the same mouse over the course of several trials comprising a single experiment for which a manually extracted ROI overlay was constructed. A time-averaged frame was constructed for this dataset which is shown behind the extracted ROIs in Figures 1 (d) and 4. Highly active axons show up as brighter then the surroundings, while dark areas correspond to areas of little or no activity. As visible in Figure 4, the larger dark

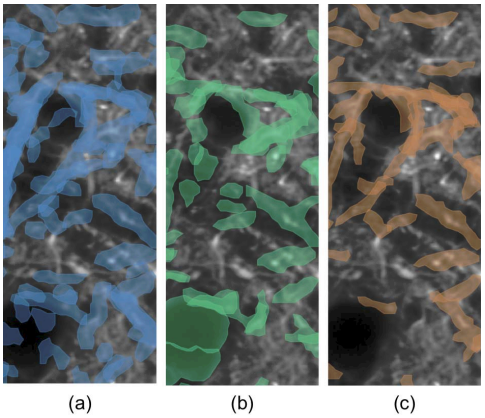

Figure 4. (a) The results of the proposed method of stICA and pixelwise classification. (b) Using only PCA instead of stICA results in a large amount of misclassification, especially in cases where there are blood vessels, the large inactive areas. (c) Using stICA, but not pixel classification finds fewer ROIs, but has a lower rate of falsely positive pixels.

areas indicate likely locations for blood vessels in which there is no neuronal activity.

The first experiment conducted was to construct an overlay of the ROIs and view it on the motion corrected version of the input video (see supplemental video). Performing this qualitative analysis of the video yielded a positive result as the obvious areas of high activity are identified and well segmented. Parameters for the minimum ROI size and μ are set to default values of 50 and 0.01 respectively. These values are found to perform well on our example 250x215 pixel datasets, but may be fine-tuned to provide a better fit for individual data.

A quantitative comparison was then performed between the automatically extracted ROIs and samples of the researcher's hand-drawn "ground truth" ROIs. In order to perform this comparison, the 30 most significant ROIs from the hand drawn set, as determined by average activity and size were lined up with the closest matching of the automatically generated ROIs. Matching was determined per significant ROI on a scale from 0 to 1 in accordance with Dice's coincidence index [3]. The Dice coefficient was then summed for all 30 pairs of ROIs in order to determine the "matching score" for each ROI extraction trial:

$$S = \sum_{j=1}^{30} \frac{2 \times |Q_j \cap D_j|}{|Q_j| + |D_j|} \tag{2}$$

Where Q_j represents the pixels in one of the significant ROIs, while D_j is the corresponding automatically generated ROI which maximizes this score. A score of 0 indicates that no overlapping automatically generated ROI is found, while a score of 1 would indicate a perfect match (see supplemental materials for examples). Summing the score for all 30 of the significant ROIs gives a score for the entire trial ranging from 0 to 30. This test was performed multiple times since the convergence of the FastICA algorithm is dependent on its random initialization. The mean score observed on the example data across 4 trials was 14.1. Examples of the extracted ROIs are shown in 4(a) and the quantitative analysis is summarized in 5.

In order to validate the importance of the stICA method, a second test was used which performed PCA followed directly by pixelwise classification. This system is similarly scored against the top 30 significant hand-drawn ROIs resulting in an average dice coefficient score of 12.9. An example of the ROIs extracted are shown in Figure 4 (b), in this case misaligned ROIs can be specifically observed over non-active blood vessel regions of the field of view.

A final test was performed in order to investigate how well stICA performed without the pixelwise classification. A portion of the resulting ROI set shown in Figure 4 (c). In order to test this the stICA components were filtered with smoothing and noise removal techniques in order to remove as much of the noise in the components as possible. These filtered components were segmented spatially and ROI boundaries were extracted. This technique tended to result in fewer misaligned ROI segments then the PCA only technique but extracted fewer total ROIs and resulted in an alignment score of 13.1, which is lower then the case of stICA followed by classification.

The results of this quantitative analysis are summarized in Figure 5. ROI extractions were completed multiple times to account for the performance due to the small deviations in convergence of the FastICA algorithm. This did lead to a single trial where both of the extraction methods relying on stICA underperformed the PCA analysis. The "Matching score" for these trials were well below the standard error and this occurrence is determined to be atypical. For the applications of ROI extraction there are significant benefits to have more ROIs extracted even at the cost of a few more false-positive areas. As long as the population of misaligned ROIs are not extremely high, researchers can manually or automatically remove them during subsequent analysis as the fluorescence traces from these areas will remain flat throughout the duration of the experiment. Missing an active region, however, comes at a higher cost since extracting neuronal activity is the main goal of this imaging.

5. Conclusion

This paper presents a novel automated procedure for axon recognition. To the best of our knowledge, the proposed method is unique and serves as an important computer vision application. A natural flow from Component extraction to ROI segmentation is discussed, and both qualitative and quantitative results are provided. The methods discussed here are effective at drawing regions of interest

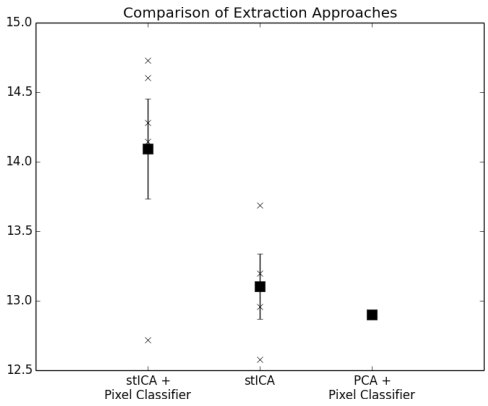

Figure 5. Extracted ROIs are assessed as being aligned or misaligned with the hand drawn "ground truth". The alignment of the 30 most significant ROIs are assessed by calculating the dice coefficient. Mean and standard error for four trials of stICA components are used due to the performance dependance on random initialization of the FastICA algorithm. For these trials $k = 40$, $\mu = 0.01$, and the patch-gradient neighborhood is 8.

around areas whose function is correlated both in space and time.

Some of the extracted ROI boundaries still appear to be overlapping or over-segmented as observed in Figure 4 (a). This problem is currently handled through automated identification and merging of the most obvious cases. Further improvements through a post-processing step involving a second classification process to reassemble and merge ROIs with nearly identical fluorescence traces are currently being investigated. Additional information, such as stICA component, ROI orientation, size and distance, may also be leveraged as descriptive features.

6. Acknowledgements

We thank members of Attila Losonczy's laboratory at Columbia University for the two-photon datasets used for this project and for advice on neuronal imaging and segmentation methods.

References

[1] A. Buades, B. Coll, and J. Morel. A non-local algorithm for image denoising. In *CVPR*, 2005. 1

[2] H. Cai, X. Xu, J. Lu, J. W. Lichtman, S. Yung, and S. T. Wong. Repulsive force based snake model to segment and track neuronal axons in 3d microscopy image stacks. *NeuroImage*, 32(4):1608–1620, 2006. 2

[3] L. R. Dice. Measures of the amount of ecologic association between species. *Ecology*, 26(3):pp. 297–302, 1945. 5

[4] D. A. Dombeck, A. N. Khabbaz, F. Collman, T. L. Adelman, and D. W. Tank. Imaging large-scale neural activity with cellular resolution in awake, mobile mice. *Neuron*, 56(1):43–57, Oct 2007. 2

[5] L. Grady. Random walks for image segmentation. *IEEE TPAMI*, 28(11):1768–1783, 2006. 1

[6] A. Hyvarinen. Fast and robust fixed-point algorithms for independent component analysis. *Neural Networks, IEEE Transactions on*, 10(3):626–634, May 1999. 3

[7] J. H. Jennings and G. D. Stuber. Tools for resolving functional activity and connectivity within intact neural circuits. *Current biology : CB*, 24(1):R41–50, Jan. 2014. 1

[8] E. Jurrus, M. Hardy, T. Tasdizen, P. T. Fletcher, P. Koshevoy, C.-B. Chien, W. Denk, and R. Whitaker. Axon tracking in serial block-face scanning electron microscopy. *Medical image analysis*, 13(1):29–39, 2008. 2

[9] P. Kaifosh, M. Lovett-Barron, G. F. Turi, T. R. Reardon, and A. Losonczy. Septo-hippocampal GABAergic signaling across multiple modalities in awake mice. *Nat Neurosci*, 16(9):1182–1184, Sept. 2013. 2

[10] E. Meijering. Cell segmentation: 50 years down the road [life sciences]. *Signal Processing Magazine, IEEE*, 29(5):140–145, 2012. 2

[11] E. A. Mukamel, A. Nimmerjahn, and M. J. Schnitzer. Automated analysis of cellular signals from large-scale calcium imaging data. *Neuron*, 63(6):747–760, 2009. 1, 2

[12] S. Osher and N. Paragios. *Geometric level set methods in imaging, vision, and graphics*. Springer, 2003. 1

[13] S. Paris and F. Durand. A fast approximation of the bilateral filter using a signal processing approach. In *ECCV*, 2006. 1

[14] S. Schmitt, J. F. Evers, C. Duch, M. Scholz, and K. Obermayer. New methods for the computer-assisted 3-d reconstruction of neurons from confocal image stacks. *Neuroimage*, 23(4):1283–1298, 2004. 2

[15] J. Shi and J. Malik. Normalized cuts and image segmentation. *IEEE TPAMI*, 22(8):888–905, 2000. 1

[16] K. Smith, A. Carleton, and V. Lepetit. General constraints for batch multiple-target tracking applied to large-scale videomicroscopy. In *CVPR*, 2008. 2

[17] J. Stone, J. Porrill, N. Porter, and I. Wilkinson. Spatiotemporal independent component analysis of event-related fmri data using skewed probability density functions. *NeuroImage*, 15(2):407–421, 2002. 1, 2

[18] L. Tian, J. Akerboom, E. R. Schreiter, and L. L. Looger. Chapter 5 - neural activity imaging with genetically encoded calcium indicators. In T. Knöpfel and E. S. Boyden, editors, *Optogenetics: Tools for Controlling and Monitoring Neuronal Activity*, volume 196 of *Progress in Brain Research*, pages 79 – 94. Elsevier, 2012. 1

[19] J. Wang, S.-F. Chang, X. Zhou, and S. T. Wong. Active microscopic cellular image annotation by superposable graph transduction with imbalanced labels. In *CVPR*, 2008. 2

[20] J. Wang, X. Zhou, P. L. Bradley, S.-F. Chang, N. Perrimon, and S. T. Wong. Cellular phenotype recognition for high-content rna interference genome-wide screening. *Journal of Biomolecular Screening*, 13(1):29–39, 2008. 2

[21] J. Wang, X. Zhou, J. Lu, J. Lichtman, S.-F. Chang, and S. T. Wong. Dynamic local tracing for 3d axon curvilinear structure detection from microscopic image stack. In *IEEE International Symposium on Biomedical Imaging*, pages 81–84, 2007. 2

[22] M. Wang, X. Zhou, F. Li, J. Huckins, R. W. King, and S. T. Wong. Novel cell segmentation and online svm for cell cycle phase identification in automated microscopy. *Bioinformatics*, 24(1):94–101, 2008. 2

[23] M.-C. Yang, C.-S. Huang, J.-H. Chen, and R.-F. Chang. Whole breast lesion detection using naive bayes classifier for portable ultrasound. *Ultrasound in medicine & biology*, 38(11):1870–1880, 2012. 2, 3